Truth approximation by concretization

Kuipers, Theo A.F.

Published in:
EPRINTS-BOOK-TITLE

Publication date:
1992

Poznań Studies in the Philosophy of the Sciences and the Humanities
1992, Vol. 25, pp. 159-179

Theo A.F. Kuipers

TRUTH APPROXIMATION BY CONCRETIZATION*

Introduction

The idea of truthlikeness or verisimilitude is that one theory can be closer or more similar to the truth than another. In 1974 Miller and Tichý proved that Popper's original definition was inadequate for it did not leave room for false theories, i.e. theories having empirical counterexamples. In Kuipers [1982, 1984, 1989] I presented and elaborated a naive structuralist definition of truthlikeness leaving room for empirical counterexamples and which was moreover attractive in other conceptual, logical and methodological respects. However, this naive definition is based on the assumption that the other kind of mistakes a theory can make, i.e. allowing mistaken models, are all equally bad. For this reason the naive definition does not seem to have real life scientific examples, for in cases of scientific progress a theory with mistaken models is usually replaced by a theory with less mistaken, but nevertheless mistaken, models. A paradigmatic case is the theory resulting from a concretization of an idealized theory.

A sophisticated definition of truthlikeness should hence not only account for empirical counterexamples but also for the fact that one mistaken model may be more similar to a required model than another. For then there is room for improving a theory by introducing new, but less, mistaken models. Of course, a sophisticated definition should reduce to the naive definition under the relevant assumptions. Finally, it should retain the attractive logical and methodological features of the naive definition.

In Kuipers [1987] I presented a first attempt at a refined definition, but for several reasons it is not satisfactory. In Kuipers [1990] a new refined definition is presented, together with a detailed analysis of the conceptual and methodological consequences of the naive and the refined definition.

The main consequences are the following. In contrast to the naive definition, the refined one is under no condition victim of Graham Oddie's child's play objection to certain definitions of truthlikeness. There are plausible, naive and refined, definitions of the claim that one theory is instantially and explanatorily more successful than another, and such success dominance can be explained by the hypothesis that the first theory is closer to the truth than the second in the naive and refined sense, respectively. On this basis it is articulated in what sense the methodological rule of success, viz. to choose the more successful theory, and related rules are functional for approaching the truth. It is also shown that the naive and the refined explanatory clause, here presented in terms of explanatory mistakes, also explicate basic intuitions concerning the explanation of laws. For the case of theories stratified with an observational and a theoretical level it is shown that, in contrast to naive truthlikeness, refined truthlikeness on the theoretical level is projected on the observational level under very interesting additional conditions. Finally it is shown that the naive definition has a plausible quantitative version, but the refined one not, for reasons which plea for the present qualitative approach.

In this article I confine myself to both definitions and the most important formal properties (Sections 1 and 3), the notion of structure-likeness underlying the refined definition (Section 2) and the application to idealization and concretization (Section 4). In this final section it is shown that 'idealization and concretization' is a special kind of potential refined truth approximation. This is illustrated by Van der Waals's theory of gases. Moreover, it is indicated how idealization and concretization can function as a strategy in validity research around 'interesting theorems'.

1. Naive truthlikeness of theories

Preparations

Let there be given a domain *D* of natural phenomena (states, situations, systems) to be investigated. *D* is supposed to be circumscribed by some informal, intensional description and may be called the primitive set of intended applications. Let there also be given a set *Mp* of *conceptual possibilities* or potential models designed to characterize *D*. It may be assumed that *Mp* is, technically speaking, a set of structures of a certain similarity type. In practice *Mp* will be the conceptual frame of a research program for *D*.

The confrontation of D with Mp, i.e. D seen throuth Mp, is assumed to generate a unique, time-independent subset $Mp(D) = T$ of all Mp-representations of the members of D, to be called the Mp-set of intended applications or the (Mp-) set of *empirical possibilities*. This assumption will be called the *frame-hypothesis* associated with $<D, Mp>$. By consequence, $Mp - T$ contains the relevant *empirical impossibilities*. As a rule, T is unknown, or even the great unknown and hence the target of theory-directed research in the domain.

It is clear that T is Mp-dependent, hence T is *conceptually relative*. It is also clear that T depends on reality through D. However, it does not represent 'the actual world', i.e. some actual state, situation or system, but the set of 'empirically possible worlds' (as far as D is concerned). For that reason, the present type of realism may be called *theoretical realism* instead of descriptive realism.

A theory is any combination of a subset X of Mp and the claim that T is equal to X, and will be briefly indicated by 'theory X' or just 'X'. Members of X are called models of theory X. Theory X is true or false when its claim '$T = X$' is true or false, respectively. According to this definition there is only one true theory, viz. theory T itself. Hence T may be called 'the true theory' or 'the theoretical truth' or even 'the truth'.

T can easily be interpreted as 'the strongest law'. A (general) hypothesis is defined as the combination of a subset X of Mp and the (weak) claim that T is a subset of X (i.e. all empirical possibilities satisfy the conditions of X). Hypothesis X is true or false when its claim '$T \subseteq X$' is true or false, respectively. Members of X are now also called models of hypothesis X. A true hypothesis is also called a law.

If Y is a subset of X the claim of hypothesis Y implies the claim of hypothesis X. In that case hypothesis Y is also said to imply hypothesis X, in agreement with standard modeltheoretic usage to say that a statement $S1$ logically implies the statement $S2$ iff the models of $S1$ form a subset of those of $S2$.

If (hypothesis) Y implies the law X it is said to explain it. Hypothesis T is of course the strongest law for it is not only true as hypothesis, i.e. it is a law, but it implies, hence explains, all other laws.

Note that the presented distinction between hypothesis and theory has nothing to do with theoretical terms. Here both a hypothesis and a theory may or may not have theoretical terms. The crucial distinction between a hypothesis and a theory in this article is that the claim of hypothesis X is just one conjunct of the combined claim of the corresponding theory X.

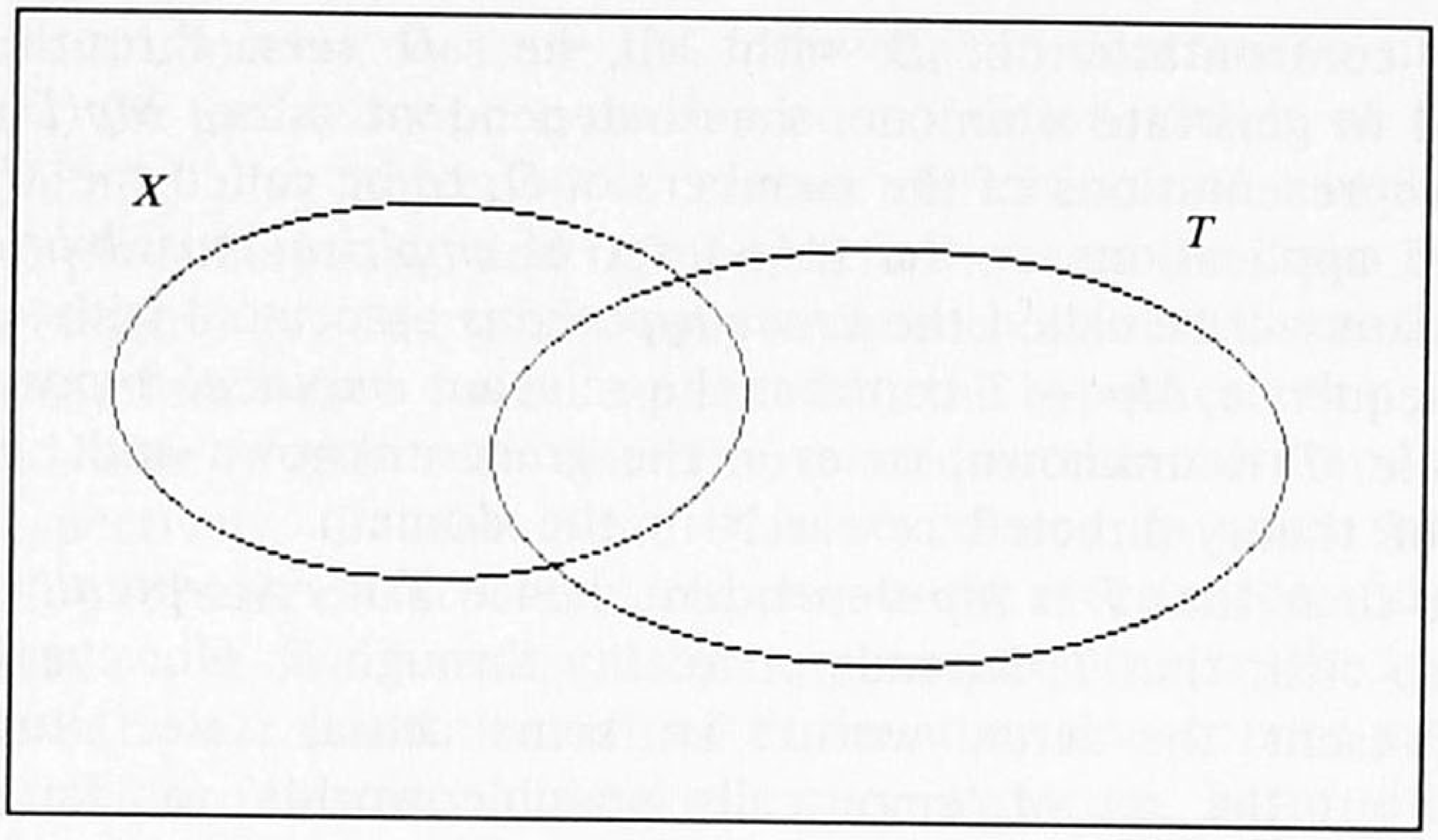

Diagram 1

Theory X can make two kinds of *mistakes*, see *Diagram 1*. The members of $T - X$, if any, are *instantial* mistakes: empirical possibilities that are excluded by X; in other words, they are the empirically realizable counterexamples of X. Recall that X explains T when T includes X. Hence, the members of $X - T$, if any, may be called the *explanatory* mistakes of X (with respect to T): empirical *im*possibilities that are not excluded by X, that is, wrongly admitted models, also called mistaken models. Note that the explanatory mistakes form, by definition, a kind of counterexamples that cannot be empirically realized. The set of all mistakes of theory X is the union of these two sets $T - X$ and $X - T$, which is technically called the symmetric difference between X and T, indicated by $X \vartriangle T$.

A theory does not only make mistakes but makes also *matches*. $T \cap X$ represents the instantial matches: empirical possibilities that are recognized as such by X or, in other words, they are the empirically realizable examples of X. Let cX indicate the complement of X with respect to Mp, i.e. $Mp - X$. X explains T, i.e. T includes X, is equivalent to cX includes cT. Hence, the members of $cT \cap cX$ are the explanatory matches: empirical *im*possibilities that are rightly excluded by X. The explanatory matches are a kind of examples that cannot be empirically realized. The union of the two sets of matches of theory X is of course equal to the complement of the total set of mistakes, viz. $c(T \vartriangle X)$.

Note that all mistakes and matches are ultimate, in the sense that they need not have been established.

Diagram 2 gives a survey of the instantial and explanatory matches and mistakes of a theory.

	matches	mistakes	total (union)
instantial	$T \cap X$	$T-X$	T
explanatory	$cT \cap cX$	$X-T$	cT
total (union)	$c(T \Delta X)$	$T \Delta X$	Mp

Diagram 2

All concepts introduced thusfar, and most of the ones to be introduced, can be illustrated by the following electric circuit (see *Diagram 3*). Let p_i for $1 \leqslant i \leqslant 4$ indicate that switch i is on (<->) and $-p_i$ that it is off (↕). Let $p_o(-p_o)$ indicate that the bulb lights (does not light). It is assumed that the bulb is not defect and that there is enough voltage. An arbitrary conceptual possibility, for instance, can be represented by a set of unnegated p_i's and the true theory about the circuit by the propositional formula p_o <--> $(((p_1 \ \& \ p_2) \vee p_3) \ \& \ p_4)$.

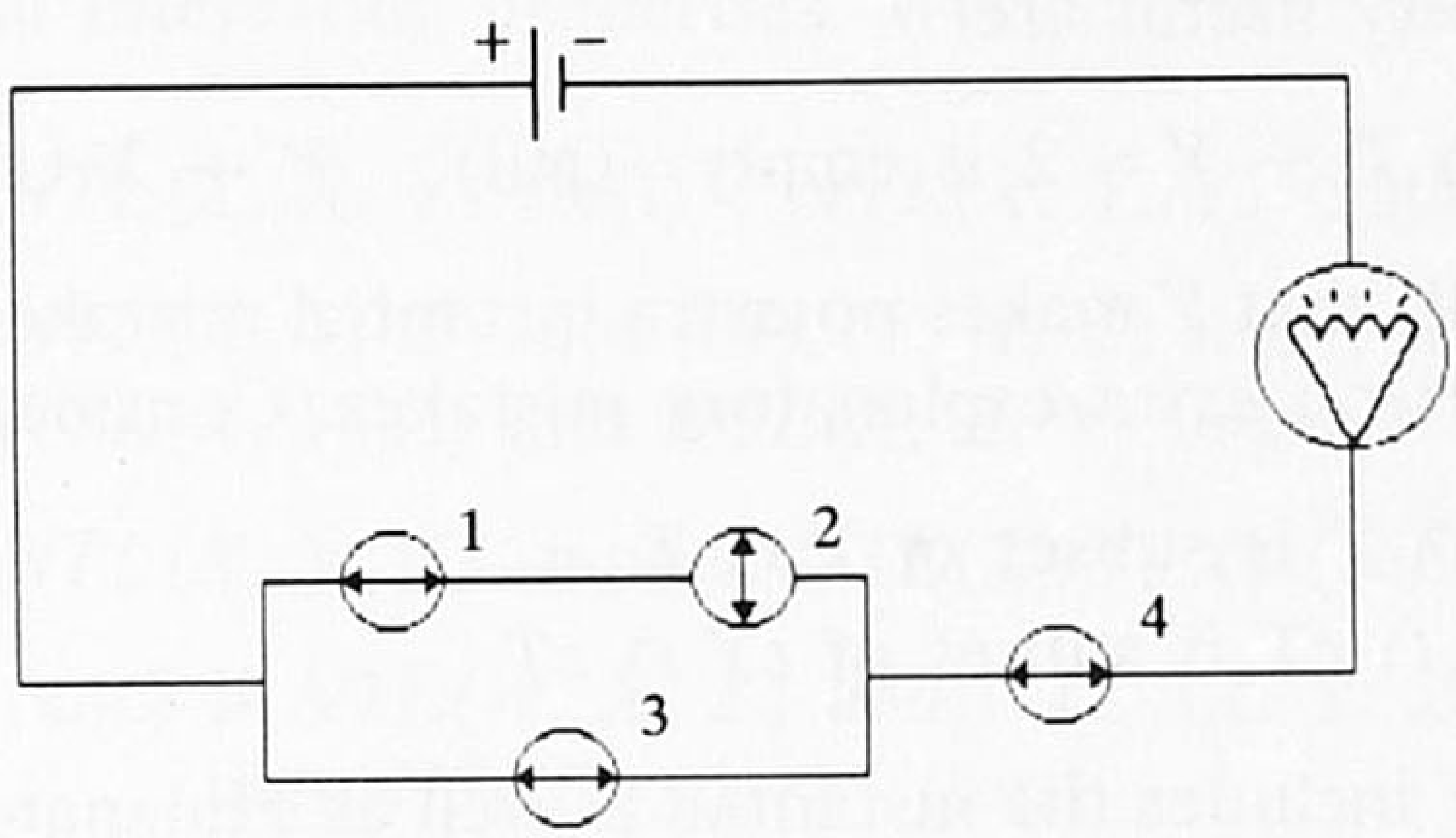

Diagram 3

The naive definition

From now on *X*, *Y*, etc. refer to theories or just to the sets *X* and *Y*, depending on the context. When the hypotheses *X* and *Y* are intended it will be explicitly mentioned.

The naive definition of truthlikeness states that *theory Y is at least as similar (close) to the truth (T) as theory X*, indicated by *NTL*(*X*, *Y*, *T*), if the following two conditions are satisfied:

(Ni) $T - Y$ is subset of $T - X$

(Nii) $Y - T$ is subset of $X - T$

The *instantial clause* (Ni) says that the instantial mistakes of X include those of Y, and the *explanatory clause* (Nii) that the explanatory mistakes of X include those of Y. Hence, it may be said that (Ni) and (Nii) require that Y instantiates and explains T at least as well as X, respectively. Note that (Nii) implies

(Nii*) when $X - T$ is empty, $Y - T$ is empty

that is the claim that Y explains T as soon as X explains T. Note also that (Ni) and (Nii) together are equivalent to the claim that the mistakes of Y ($Y \vartriangle T$) form a subset of those of X ($X \vartriangle T$).

By NTL + (X, Y, T) I indicate that Y is more similar to T than X in the strict sense that the mistakes of Y form a *proper* subset of those of X. Here and later the strong verbal expressions 'closer to' or 'more similar to' will however also be used to refer to the corresponding weak notion. When the strict notion is meant it will be explicitly stated.

Some equivalent formulations of *NTL* are instructive. The numbers of the sets refer to *Diagram 4* (in which *Mp* is not explicitly indicated). Practically very useful are:

(Ni)' $X \cap T - Y = 2$ is empty (Nii)' $Y - X \cup T = 6$ is empty

The first tells that Y makes no extra instantial mistakes, and the second that Y makes no extra explanatory mistakes. Consider also:

(Ni)'' $X \cap T$ is subset of $Y \cap T$
(Nii)'' $cX \cap cT$ is subset of $cY \cap cT$

telling that Y includes the instantial as well as explanatory matches of X, respectively.

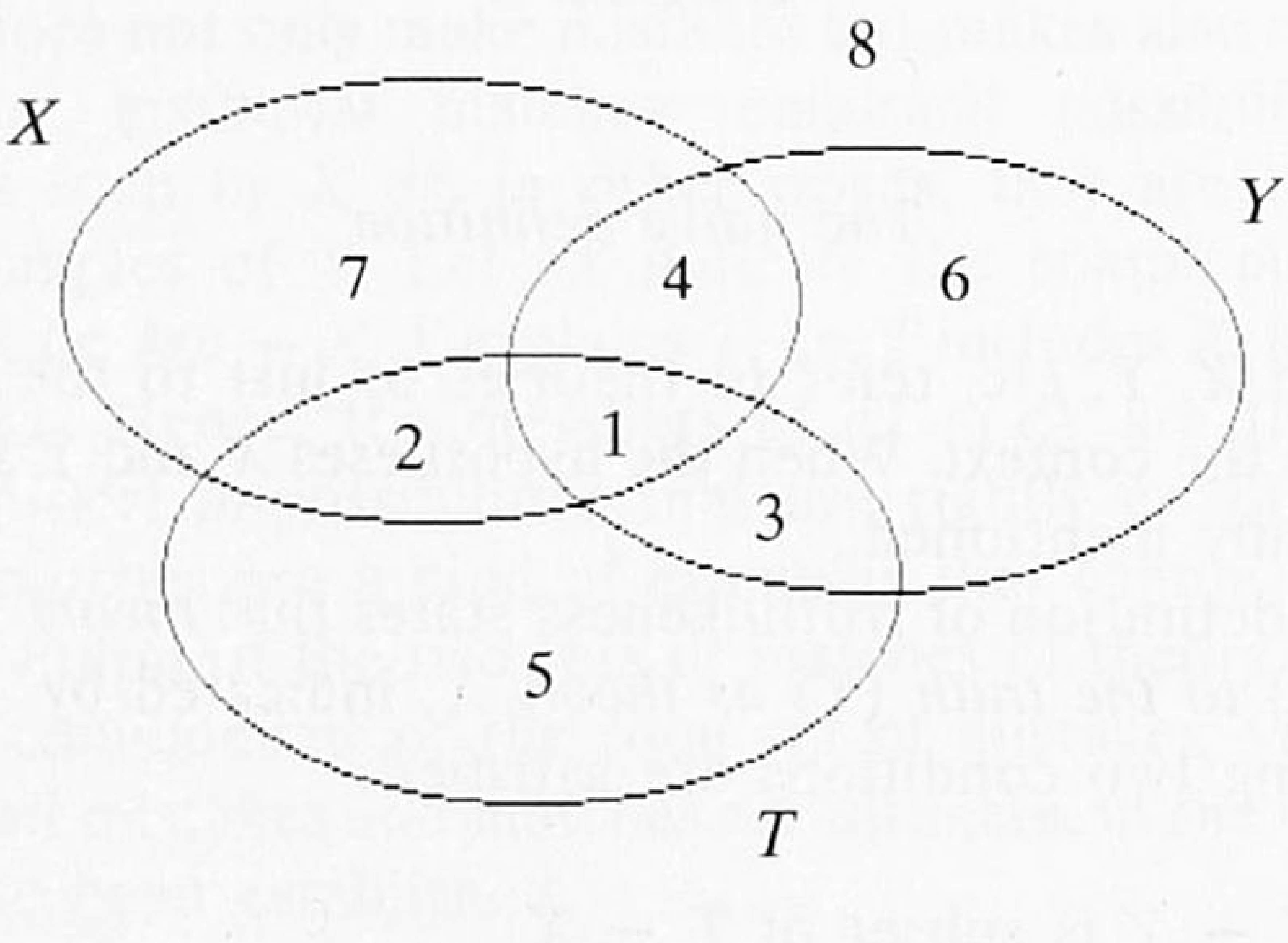

Diagram 4

It is important to note that improving a theory X in the sense of finding a theory Y such that $NTL+(X, Y, T)$ is not an easy task, due to the fact that both components are counteracting. This can nicely be illustrated by considering e.g. just weakening of theory X: if Y is weaker than X (i.e. $Y \supseteq X$) then Y instantiates T as well as X, but X explains T as least as well as Y. Of course, strengthening a theory leads to the opposite tension.

Some formal properties

Any definition of the binary relation of truthlikeness between theories X and Y can be seen as a special case of a ternary relation of *theorylikeness* $NTL(X, Y, Z)$ between theories X, Y and Z by replacing the fixed true theory T by the variable theory Z. Then we get the general definition: $NTL(X, Y, Z) =_{df} Z - Y \subseteq Z - X$ and $Y - Z \subseteq X - Z$, with plausible equivalent formulations.

NTL has several interesting properties. We list the main ones.

reflexivity:	$NTL(X, X, Y)$ (left) $NTL(X, Y, Y)$ (right)
antisymmetry:	$NTL(X, Y, Z)$ and $NTL(Y, X, Z)$ imply $X = Y$ (left)
	$NTL(X, Y, Z)$ and $NTL(X, Z, Y)$ imply $Y = Z$ (right)
symmetry:	$NTL(X, Y, Z)$ implies $NTL(Z, Y, X)$ (central)
transitivity:	(e.g.) if $NTL(W, X, Z)$ and $NTL(X, Y, Z)$ then $NTL(W, Y, Z)$ (left)

Hence, from left reflexivity, left antisymmetry and left transitivity it follows that $NTL(X, Y, Z)$ is for fixed Z a partial ordering of theories. By consequence, a sequence of theories converging to the truth is perfectly possible.

Some other interesting properties are:

centeredness:	$NTL(X, X, X)$
centering:	if $NTL(X, Y, X)$ then $X = Y$
specularity:	if $NTL(X, Y, Z)$ then $NTL(cX, cY, cZ)$
concentricity:	if $X \subseteq Y \subseteq Z$ then $NTL(X, Y, Z)$ and $NTL(Z, Y, X)$
context neutrality:	if X, Y and Z are subsets of Mp and Mp itself a subset of a larger set of conceptual possibilities Mp^* then $NTL(X, Y, Z)$ implies $NTL^*(X, Y, Z)$

2. Structurelikeness and truthlikeness of structures

Up to now I was dealing with the problem of truthlikeness of theories and more generally theorylikeness. But there is also a problem of truthlikeness of structures and more generally structurelikeness. In the circuit example (*Diagram 3*) for instance it is clear that there is, given the conceptual frame, not only just one true theory characterizing the set of empirically possible states of that particular circuit. There is also just one true description of the actual state of the circuit as it is depicted, p_0 & p_1 & p_2 & p_3 & p_4, according to the standard propositional representation. In general, in addition to the frame-hypothesis leading to the assumption that there is just *one true theory*, I will assume that, given a conceptual frame, every particular situation or state of affairs (of a system) in the domain, every actual world so to speak, has just one correct representation or *one true description*. By consequence, with each experiment, i.e. with each realization of an empirical possibility, there is associated a unique true description within the conceptual frame.

Hence, the traditional problem of explicating the idea of truthlikeness concerns on closer inspection two intuitive phrases, viz. "one description is more similar to the true theory than another". In the previous section I dealt with the second problem, neglecting the fact that it is plausible to take into account a possible underlying notion of likeness of descriptions. This will be done in the next section. In the present section I will just deal with the idea of truthlikeness of descriptions or, equivalently from the structuralist point of view, truthlikeness of structures and more generally structurelikeness.

Let x, y, z indicate structures in *Mp* and $s(x, y, z)$ indicate that y is at least as similar (close) to z as x. The true structure of the context will be indicated by t. I will not aim at a general definition of structurelikeness, for a precise definition will have to depend on the specific nature of the conceptual possibilities.

When $s(x, y, z)$ y is said to be *between*, or an *intermediate* of, x and z. Structurelikeness is not generally assumed to be symmetric: $s(x, y, z)$ does not generally imply $s(z, y, x)$. By consequence, being in between or an intermediate may be a directed notion: if y is between x and z in the sense of $s(x, y, z)$, this does not yet imply that y is also between z and x in that sense of structurelikeness.

Structures x and z are said to be connected or *related*, $r(x, z)$, if and only if there is y such that $s(x, y, z)$. It follows that r is not by definition symmetric, i.e. $r(x, z)$ does not automatically imply $r(z, x)$. But $r(x, z)$ is already guaranteed by $s(x, x, z)$ or $s(x, z, z)$. Hence, the basic idea behind $r(x, z)$ is not the existence of a *proper* intermediate, but only that x and z

have at least so much in common that it makes sense to talk about (proper and improper) intermediates: in other words, they may also be said to be *comparable*. For instance, all pairs of propositional structures, see below, as in the case of the circuit example, will be comparable if *Mp* contains only structures constituted by one set of elementary propositions. But as soon as structures based on subsets of this set are also taken into consideration not all pairs are comparable anymore. p & q and $-p$ & $-q$ have an intermediate, e.g. p & $-q$, but p & q and p don't have. The case of concretization below (Section 4) provides another example of this situation, e.g. not every Van der Waals gas model is a concretization of every ideal gas model (they may deal with different sets of states and/or different numbers of moles). Hence, they don't need to have an intermediate. In general, a minimal condition for comparability seems to be that the two structures have the same base- or domain-sets.

It would have been possible to introduce $r(x, z)$ as a primitive term. But I have not done so because it always seems possible to read $r(x, z)$ directly off from the relevant specific definition of $s(x, y, z)$.

In the next section we will see that naive truthlikeness as defined in the previous section is in fact based on *trivial* structurelikeness, indicated by $t(x, y, z)$, defined by $x = y = z$. (The symbol 't' is used in this text as a time variable, and to indicate trivial structurelikeness and 'the descriptive truth', but confusions need not arise). Note that in the case of trivial structurelikeness two structures are only related when they are equal. Hence, naive truthlikeness is based on the idea that two different structures are never comparable.

Properties

Let us assume some plausible properties of the notion of structurelikeness. s is *centered* iff $s(x, y, x)$ implies $x = y$. s is said to be *conditionally left/right reflexive* iff $s(x, y, z)$ implies all kinds of left and right reflexivity, i.e. $s(x, x, y)$, $s(x, x, z)$, $s(y, y, z)$ and $s(x, y, y)$, $s(x, z, z)$, $s(y, z, z)$, respectively. Note that $r(x, z)$ now implies $s(x, x, z)$ and $s(x, z, z)$. Together these properties are called *the minimal s-conditions*. Note that being centered implies that r is reflexive, but the converse does not hold.

s is called *symmetric* when $s(x, y, z)$ implies $s(z, y, x)$ and *antisymmetric* when $s(x, y, z)$ and $s(z, y, x)$ imply $x = y = z$, in which case centering trivially follows. If s is symmetric then r is symmetric as well, and if s is antisymmetric then r is antisymmetric. Note that the converse implications do not hold.

There are many ways in which s can be *transitive*, e.g. left transitivity: $s(w, x, z)$ and $s(x, y, z)$ imply $s(w, y, z)$. However, none of these ways implies that r is transitive, as a laborious survey makes clear, nor does the transitivity of r imply any of these ways. Moreover, and this is even more important to note, r can be transitive without assuming that the middle term is the intermediate: if $r(x, y)$ and $r(y, z)$ then $r(x, z)$ may generally be the case, without implying that $r(x, z)$ is due to $s(x, y, z)$.

As far as r is concerned, it is useful to state in sum that r may well be an equivalence relation or a partial ordering, without strong implications for s. In the case of an equivalence relation comparability generates equivalence classes of comparable structures. In the case of a partial ordering directed sequences of comparable structures arise.

In Kuipers [1990] a number of examples of symmetric structurelikeness are presented, viz. concerning propositional, first order, and real number structures. In Section 4 we will see that concretization provides a good example of antisymmetric structurelikeness, generating a partial ordering of comparable structures.

3. Refined truthlikeness of theories

It is clear that the naive definition of truthlikeness of theories does not exploit the idea that one structure may be more similar to a second than a third, i.e. the idea of an underlying notion of structurelikeness. Let us assume that there is such an underlying ternary relation of structurelikeness s and that it satisfies the minimal s-conditions introduced in Section 2: being centered, centering (together: $s(x, y, x)$ iff $x = y$) and conditional left and right reflexivity ($s(x, y, z)$ implies e.g. $s(x, x, y)$ and $s(y, z, z)$).

I will present a refined definition of truthlikeness of theories which turns out to have many plausible and desirable properties if we neglect a queer kind of theories and restrict our attention to so-called convex theories. A set X is called *convex* (with respect to s) if it is closed for intermediates, i.e. if for all x and z in X and all y if $s(x, y, z)$ then y is in X.

Note that there are already elementary examples of non-convex theories. Assume that the definition of propositional structurelikeness is such that "p & $-q$" is between "p & q" and "$-p$ & $-q$". Then it is for instance easy to see that the propositional theory indicated by "p<-->q" is non-convex, for the non-model "p & $-q$" is between the models "p & q" and "$-p$ & $-q$" of that theory. However, it is doubtful whether theories which are non-convex, with respect to the relevant underlying notion of structurelikeness, play an important role in science proper. If, for instance, T is not convex, this has the consequence that there is an empirical *im*possibility (in the relevant sense) between two empirical possibilities, and this is unlikely as far as nature is continuous. Moreover, in Section 4 we will see that there is an additional reason why theories are convex with respect to structurelikeness based on concretization.

However this may be, many of the results to be presented are simply invalid or need technical qualifications if the relevant local convexity assumption is not satisfied, but I will not specify such qualifications. Instead of just dealing throughout with convex theories or specifying the situation for non-convex cases, I have chosen for the middle course of locally indicating when convexity is a necessary condition for the reported result. The first option can be obtained from the present text by just skipping all local convexity assumptions and assuming convex theories throughout.

I will introduce the refined definition of "Y is at least as similar to T as X" again by an instantial and an explanatory clause.

The *refined instantial clause* expresses again the intuitive idea that Y instantiates T at least as well as X, but now not just in the sense of full instantial matches, but also in the sense of approximate matches:

(Ri) for all x in X and z in T
if $r(x, z)$ then there is y in Y such that $s(x, y, z)$

It is easy to check that (Ri) implies the corresponding naive instantial clause (Ni) because s satisfies centering. Hence, it is a strengthening of the naive clause. A reformulation of (Ri) is instructive. Due to conditional reflexivity (Ri) is equivalent to:

for all x in $X - Y$ and z in $T - Y$
if $r(x, z)$ then there is y in Y such that $s(x, y, z)$

that is, for every extra model of X comparable to an instantial mistake of Y, Y has a model which is at least as similar to that mistake.

The *refined explanatory clause* will not be a strengthening but a weakening of the corresponding naive one, which required that $Y - X \cup T$ was empty. In the context of structurelikeness it is plausible to leave room for members of $Y - X \cup T$, i.e. extra explanatory mistakes of Y, provided they are between $X - T$ and T, that is, all y in $Y - X \cup T$ have to be between a member of $X - T$ and one of T. This clause guarantees as it were that Y is moving up from X to T, without detour. This results in the following clause:

(Rii) for all y in $Y - X \cup T$
there are x in $X - T$ and z in T such that $s(x, y, z)$

It is evident that the naive explanatory clause implies the refined one. (Rii) expresses the idea that every extra explanatory mistake of Y is at least as similar to some empirical possibility than some explanatory mistake of X. This may also be paraphrased by the claim that every extra explanatory mistake of Y guarantees the existence of some explanatory mistake of X that is at least as serious in the weak sense that the former is at least as similar to some empirical mistake than the latter. Hence, (Rii) expresses in this weak refined sense that Y explains T at least as well as X.

It would be possible to strengthen (Rii) by adding "and there is no z' in T such that $s(y, x, z')$" which leads to a strong refined sense of the idea that Y explains T at least as well as X. However, this did not lead to further conceptual elucidation, nor to elegant results. Given the fact that I also like to keep the explication as weak as possible, the present version of the refined explanatory clause is the most attractive one.

(Rii) is the general formulation, also appropriate for non-convex theories, in the sense of leading to centering also for such theories, see below. When T is convex (Rii) can be simplified to:

$$\begin{array}{l}\text{for all } y \text{ in } Y - X \cup T \\ \text{there are } x \text{ in } X \text{ and } z \text{ in } T \text{ such that } s(x, y, z)\end{array}$$

for the convexity of T assures that the guaranteed x in X is in $X - T$. Hence, in this case (Rii) comes down to the claim that (every member of) $Y - X \cup T$ is between (a member of) X and (a member of) T. The inclusion of non-convex cases precludes this otherwise highly plausible conceptual justification of (Rii).

The resulting definition of refined truthlikeness, i.e. *Y is at least as similar (close) to T as X*, indicated by $RTL(X, Y, T)$, imposes both the clauses (Ri) and (Rii), and may be paraphrased by: Y instantiates and explains T at least as well as X.

It is easy to state and prove the desirable reducibility of the refined to the naive definition. If s is just trivial structurelikeness t, which was defined by: $t(x, y, z)$ iff $x = y = z$, refined truthlikeness reduces to naive truthlikeness. That is, indicating $RTL(X, Y, T)$ based on $s = t$ by $RTL_t(X, Y, T)$, it is easy to prove the following *reduction theorem*: $NTL(X, Y, T)$ if and only if $RTL_t(X, Y, T)$. That (Ri) reduces to (Ni) for $s = t$ follows immediately from the fact that it implies this, as we have seen already, and that the condition is vacuous for different members of X and T, for $r_t(x, z)$ implies $x = z$, i.e. different structures are never comparable on the basis of trivial structurelikeness. On the other hand, (Rii) reduces trivially to (Nii) for $s = t$, for in that case there cannot be a member of Y outside $X \cup T$ between two different members of $X \cup T$. Note in passing that all sets are trivially convex with respect to t.

As already hinted upon for the explanatory clause, the definition of refined truthlikeness might be sharpened by strengthening one or both clauses, but I do not see convincing reasons to do so. Moreover, it is important to note that $RTL(X, Y, T)$ is a general definition in the sense that it is not based on a particular specification of structurelikeness. Any specification which is appropriate for the particular type of structures of the context leads to the relevant specific form of RTL for that context. It might be such that such a specification makes also some sharpening of the clauses plausible.

Formal properties

For the formal properties to be considered I jump to the general definition of refined *theory*likeness which is obtained from the formal version of the refined definition of *truth*likeness by replacing the true theory T by the arbitrary theory Z. In *Diagram 4* T is also supposed to be replaced by Z. It will turn out that RTL satisfies almost all properties which have been listed as properties of NTL, with some qualifications, in particular for symmetry.

Illuminating is the *sufficient condition property* (SC-property): if sets $X \cap Z - Y = 2$ and $Y - X \cup Z = 6$ and $Z - X \cup Y = 5$ and/or $X - Y \cup Z = 7$ are empty then $RTL(X, Y, Z)$. From this property immediately follow the following properties. *Concentricity*: if X is a subset of Y and Y of Z or if Z is a subset of Y and Y of X then $RTL(X, Y, Z)$, with the immediate consequence that RTL is *centered*, i.e. $RTL(X, X, X)$. Moreover, concentricity implies (unconditional) *left and right reflexivity*: $RTL(X, X, Y)$ and $RTL(X, Y, Y)$, respectively. Hence, all theories are comparable, in the sense that for all X and Z there is Y such that $RTL(X, Y, Z)$. It is also easy to prove that RTL satisfies *centering*, i.e. if $RTL(X, Y, X)$ then $X = Y$. By consequence, RTL satisfies, like NTL, the three properties that were called the minimal s-conditions for the underlying notion of structurelikeness. That these likeness-notions share these minimal formal properties is plausible and desirable: theorylikeness may well function as structurelikeness for likeness of higher order theories: sets of theories of the present exposition. Note in passing that centering would only follow for convex X, if I would not have required in (Rii) that x is an explanatory mistake of X, but simply that it is a model of X.

Another interesting property RTL shares with NTL is *context neutrality*. Let X, Y and Z be subsets of Mp. If Mp itself is a subset of an extended set of conceptual possibilities Mp^* and if s^* is an extension of s (with r^* and T^* based on s^*) then $RTL(X, Y, Z)$ iff $RTL^*(X, Y, Z)$.

The proof uses the fact that conditional reflexivity of s already guarantees for all x and z in Mp that $r(x, z)$ iff $r^*(x, z)$. Hence, theorylikeness is not disturbed by enlarging or diminution of the set of conceptual possibilities, as long as the theories themselves are not changed.

Now I turn the attention to *(anti-)symmetry*, first the central versions. Whereas naive theorylikeness was trivially symmetric, it now depens on the specific nature of the underlying notion of structurelikeness whether the ternary relation of theorylikeness is symmetric in the sense that $RTL(X, Y, Z)$ implies $RTL(Z, Y, X)$, or not. If it is not symmetric it may be antisymmetric: if $RTL(X, Y, Z)$ and $RTL(Z, Y, X)$ then $X = Y = Z$, in which case centering ($RTL(X, Y, X)$ implies $X = Y$) immediately follows.

It is easy to check that $RTL(X, Y, Z)$ is symmetric when it is based on symmetric structurelikeness, provided both X and Z are convex. In other words, the refined definition guarantees symmetry transport from the level of structures to the level of convex theories. However, antisymmetry transport from the level of structures to that of theories is not guaranteed by the refined definition. In Section 4 we will see that theorylikeness based on (antisymmetric) concretization of structures provides an antisymmetric example.

Turning to non-central symmetry notions, left antisymmetry is, in view of the possibility of sequences of theories straightforwardly converging to the truth, the most interesting notion. Under certain conditions it is not difficult to prove that left antisymmetry is transported from structurelikeness to theorylikeness: to be precise, if structurelikeness is left antisymmetric and if it is *(de-)composable*, defined by: $s(x, y, z)$ iff $r(x, y)$ and $r(y, z)$, and if X and Y are convex then $\mathrm{RTL}(X, Y, Z)$ and $RTL(Y, X, Z)$ imply $X = Y$. That s is *de*composable in the sense that $s(x, y, z)$ implies $r(x, y)$ and $r(y, z)$ follows immediately from conditional reflexivity and the definition of r. That s is *composable* in the sense that $r(x, y)$ and $r(y, z)$ together imply $s(x, y, z)$ is a substantial condition. But, as we will see in Section 4, it is trivially satisfied by the ternary relation of concretization.

NTL satisfied all kinds of *transitivity*, of which, again in view of truth convergent sequences, left transitivity is the most important one. It is not difficult to prove that left transitivity of structurelikeness is transported to refined theorylikeness, that is, it guarantees: $RTL(W, X, Z)$ and $RTL(X, Y, Z)$ imply $RTL(W, Y, Z)$. Other, but similar, transitivity results follow easily. Combining the results about left reflexivity, left antisymmetry and left transitivity we get that $RTL(X, Y, Z)$ is a partial ordering of convex theories for fixed Z if $s(x, y, z)$ is (de-)composable and a partial ordering for fixed z. Hence, under these conditions a sequence of convex theories converging to the truth is perfectly possible.

There is one property of NTL which is not at all shared by RTL, viz. *specularity*: $RTL(X, Y, Z)$ does not generally imply $RTL(cX, cY, cZ)$. This is directly related to the fact that NTL deals with instantial and explanatory mistakes in essentially the same way, whereas RTL introduces a basic asymmetry between these kinds of mistakes. The first proposal for a definition of refined truthlikeness [Kuipers 1987] struck to the symmetric treatment of the two kinds of mistakes. It came down to (Ri) and as second clause (Ri) applied to the complements of X, Y and T/Z. A number of objections raised by Van Benthem [1987] to NTL and that refined proposal were essentially due to the symmetric treatment of instantial and explanatory mistakes, which prevented for instance the allowance of extra explanatory mistakes for the better theory.

4. Application: idealization & concretization

In this section I will study a special kind of theorylikeness, viz. theorylikeness based on idealization and concretization. From the general exposition it then trivially follows that concretization of theories can be a truth approximation strategy. This will be illustrated by the transition of the theory of ideal gases to that of Van der Waals. Then I will outline how concretization is also an important strategy in the investigation of the domain of validity of an interesting theorem and in particular whether it is true for the actual or even the empirically possible worlds.

Idealization and concretization

Concretization or factualization, as it has been presented by the Polish philosophers Władysław Krajewski [1977] and Leszek Nowak [1980], is basically a relation between real-valued functions. Hence, let us assume that the conceptual structures to be considered contain one or more real-valued functions, with or without one or more real constants. Structure y is called a *concretization* of x and x an *idealization* of y, indicated by $con(x, y)$, if y transforms, directly or by a limit procedure, into x when one or more constants (functions) occurring in y (uniformly) assume the value 0. It is easy to see that it is a necessary condition for $con(x, y)$ that x and y have the same domain-sets. Moreover it is easy to check that con is reflexive, antisymmetric and transitive. In a subsection to follow the example is presented of a Van der Waals gas model as a concretization of an ideal gas model.

Concretization is primarily a binary relation, but for my purposes I need the plausible ternary version leading to a *concretization triple*:

$ct(x, y, z)$ if and only if $con(x, y)$ and $con(y, z)$. I will assume *ct* as the underlying notion of structurelikeness. The relation of relatedness based on *ct* is easily seen to be equivalent to *con*. Note that we have here a clear example in which relatedness is not symmetric, but directed. Note also that *ct* is trivially decomposable. Theory-/truth-likeness based on this ternary relation will be indicated by $RTL_{ct}(X, Y, Z)$ and $RTL_{ct}(X, Y, T)$, respectively. It is easy to check that *ct* satisfies the minimum conditions of being centered, centering and conditional left and right reflexivity. Moreover, it is antisymmetric (central, left and right) and it satisfies all conceivable kinds of transitivity, e.g. left: if $ct(w, x, z)$ and $ct(x, y, z)$ then $ct(w, y, z)$.

My next task is to define the binary relation of concretization between theories. Again I will do this as weak as possible: Y is a *concretization* of X and X an *idealization* of Y, indicated by $CON(X, Y)$, if and only if all members of X have a concretization in Y and all members of Y have an idealization in X. At first sight one might think that the second clause should be strengthened to: and all members of Y have a *unique* idealization in X. However, this would exclude e.g. 'inclusive' concretization triples $<X, Y, Z>$ with X as subset of Y and Y of Z and $CON(X, Y)$ and $CON(Y, Z)$.

It is trivial that *CON* is reflexive and transitive. However, it need not be antisymmetric as one might expect. But sufficient for antisymmetry of $CON(X, Y)$ is that X and Y are convex (i.e. closed for intermediates). Now it comes down to: if $con(x, y)$ and $con(y, z)$, i.e. $ct(x, y, z)$, and x and z in X then y in X. Note that it is in the present context highly plausible to assume that the true theory T is convex, and hence that all theories to be considered are convex. For if T would be non-convex it would imply that some idealization of an empirical possibility is empirically *im*possible, but a stronger idealization is again empirically possible. Even in cases where nature is not continuous this is not very plausible.

The ternary relation of concretization of theories I define again as weak as possible: $CT(X, Y, Z)$ if and only if $CON(X, Y)$ and $CON(Y, Z)$. It is easy to check that *CT* has the properties of being centered, centering for convex sets, conditional left and right reflexivity, antisymmetry (central, left, right) for convex sets and all conceivable forms of transitivity. By consequence, $CT(X, Y, Z)$ is for fixed Z a partial ordering as far as convex theories are concerned.

The main question is whether or under what conditions $CT(X, Y, Z)$ implies $RTL_{ct}(X, Y, Z)$. It turns out that some conditions have to be added to guarantee this implication, but there are some alternative possibilities. I am of course primarily interested in conditions on X and/or Y or their combination, for in the crucial case we do not dispose of Z, i.e. T.

One sufficient combination of conditions is the following: Y should be convex as well as *mediating*, the latter condition being defined as: if z is a concretization of x and if x has a concretization in Y and z an idealization in Y then Y provides also an intermediate for x and z, or, more formally, if $con(x, z)$ and if there are y and y' in Y such that $con(x, y)$ and $con(y', z)$ then there is y'' in Y such that $con(x, y'')$ and $con(y'', z)$ (i.e. $ct(x, y'', z)$).

Note that both conditions only concern Y. Although being mediating is a more specific property than convexity, it is not a very restrictive condition in the present context. Note also that it follows that any X can be an idealized starting point for successive concretization. However, the starting point X will usually even be *closed for idealizations* in the sense that if x in X and $con(x', x)$ then x' in X. It is easy to check that this trivially implies that X is convex and mediating.

Let us formally state the main claim: it is (easily) possible to prove the following *Concretization as Theorylikeness Theorem* (*C-->TL*-Theorem): if $CT(X, Y, Z)$ and if Y is convex and mediating then $RTL_{ct}(X, Y, Z)$. In words: the intermediate theory of a concretization triple is closer to the third than the first, assuming that it is convex and mediating.

We may define stronger versions of concretization triples such as $CT^*(X, Y, Z) = CT(X, Y, Z)$ and Y convex and mediating or even $CT^{**}(X, Y, Z) = CT^*(X, Y, Z)$ and X and Z also convex. According to the theorem both are special kinds of theorylikeness. Moreover, it was already mentioned that $CT(X, Y, Z)$ is antisymmetric (in the central sense) as soon as the three sets are convex; hence CT^{**} is an antisymmetric special type of theorylikeness.

Truth approximation

A direct consequence of the *C-->TL*-Theorem is that, if theory Y is a concretization of theory X, if Y convex and mediating, and if the true set of empirical possibilities T is a concretization of Y then Y is closer to the truth than X. This may be called the *Truth Approximation by Concretization (TAC-)Corollary*, a major goal of this section, viz. to show that and in what sense concretization may be a form of truth approximation. All conditions for truth approximation can be checked, except of course the crucial heuristic hypothesis that T is a concretization of Y.

Application to gas models

The transition from the theory of ideal gases to Van der Waals's theory of gases has frequently been presented as a paradigmatic case of concretization. The challenge of any sophisticated theory of truthlikeness hence is to show that this transition can be a case of truth approximation.

For this purpose I start with formulating the relevant models in elementary structuralist terms. $<S, n, P, V, T>$ is a *potential gas model* (*PGM*) iff S represents a set of thermal states of n moles of a gas and P, V and T are real-valued functions defined on S representing pressure, volume and (empirical absolute) temperature, respectively.

Specific gas models are *PGM*'s satisfying an additional condition. The *ideal gas models* (*IGM*) satisfy in addition $P(s)V(s) = nRT(s)$ for all s in S, or simply $PV = nRT$, where R is the so-called ideal gas constant. For *gas models with mutual attraction* (*GMa*) there is a non-negative real (number) constant a, within a certain fixed interval, such that $(P + (n^2a/V^2))V = nRT$. For *gas models with non-zero volume of molecules* (*GMb*) there is a non-negative real constant b, within a certain fixed interval, such that $P(V - nb) = nRT$. Finally, in the case of *Van der Waals gas models* (*WGM*) there are non-negative real constants a and b, within the previously mentioned two intervals, such that $(P + (n^2a/V^2))(V - nb) = nRT$.

Note first that it is a necessary condition for $con(x, y)$ (x and y in *PGM*) that $S_x = S_y$. Note also that *IGM*/*GMa*/*GMb*/*WGM* have been defined such that they are all convex and mediating.

It is easy to check that *IGM*, *GMa* and *WGM* as well as *IGM*, *GMb* and *WGM* constitute a concretization triple: an element of *WGM* transforms into an element of *GMa*/*GMb* by substituting the value 0 for b and a, respectively. The resulting elements of *GMa* and *GMb* transform into elements of *IGM* by substituting 0 for a and b, respectively.

Due to the *C-->TL*-Theorem it follows that *GMa* and *GMb* are both closer to *WGM* than *IGM*. By consequence, if *WGM* would represent the true set of empirically possible gases, *GMa* and *GMb* would be closer to the truth than *IGM*. Finally, and most importantly, the *TAC*-Corollary guarantees that *WGM* is closer to the truth than *IGM*, assuming the heuristic hypothesis that the true set of empirically possible gases is on its turn a concretization of *WGM*.

See Niiniluoto [1986] for a completely different quantitative approach to concretization in general and the Van der Waals case in particular.

Validity research

Scientific research is not always directed to describing the actual world or characterizing the set of empirically possible worlds. It may also primarily aim at proving interesting theorems for certain conceptual possibilities, as Hamminga [1983] showed for neo-classical economics.

Let a certain *Mp* have been chosen, let T indicate the (unknown) subset of empirical possibilities, and let R indicate the (unknown) subset (of T) of realized empirical possibilities, possibly containing just one element, the actual possibility.

Let *IT* indicate an 'interesting theorem', that is some insightful claim, of which it is interesting to know whether it is true for the empirical possibilities, or at least the realized possibilities. Let $V(IT)$, or simply V, indicate the set of conceptual possibilities for which *IT* can be proved. V is called the *domain of (provable) validity* of *IT*, and it is assumed to be not yet explicitly characterized.

A frequent type of scientific progress is the following. Suppose that it was earlier proved that *IT* holds for X, i.e. that X is subset of V. The new result is that Y, which includes X, also is, like X, included in V. Due to concentricity of the naive and refined theorylikeness notions it follows in this case that $N/RTL(X, Y, V)$. The ultimate purpose of this type of research was to find out whether T or at least R are subsets of V. Of course, the larger V has been proven to be, as in the described case, the greater the chance, informally speaking, that R or even T are subsets of V. However, just enlarging the proven domain of validity does not necessarily go in the direction of R and T. For this purpose concretization is the standard strategy.

Let it first have been shown that X is a subset of V, and later also that a concretization Y of X ($CON(X, Y)$, X need not be a subset of Y) is a subset of V. It then trivially follows that $RTL(X, X \cup Y, V)$. If, moreover, Y is convex and mediating, it follows from the heuristic hypothesis that T is concretization of Y ($CON(Y, T)$), using the *C-->TL*-theorem, that $RTL(X, Y, T)$. Hence, we have proved *IT* for a set Y which is more similar to T than X, which increases the chance that *IT* holds for T, *ipso facto* for R.

A complex form of validity research concerns the case that *IT* is not fixed, but that realistic factors are successively accounted for. Formally this is also a form of concretization. *IT*2 is called a concretization of *IT*1 if $V(IT2) = V2$ is a concretization of $V(IT1) = V1$.

Now suppose that *IT*1 is proved for X. The relevant heuristic strategy is to look for a concretization Y of X and a concretization *IT*2 of *IT*1 such that *IT*2 can be proved for Y. The heuristic hypotheses are that T is a concretization of Y and that there is a concretization IT^* of *IT*2 such that IT^* holds for T and hence for R. This makes sense because if Y and *IT*2 are convex and mediating it not only follows that Y is closer to T than X but also that $V2$ is closer to V^* than $V1$. Hence, in this case we are not only on the way to T but also to IT^*.

Concluding remarks

In this article I have first presented conceptually plausible definitions of naive and refined truthlikeness of theories, the latter based on an underlying notion of structurelikeness. Then the important point was shown that refined truthlikeness makes sophisticated truth approximation possible. In particular, it was pointed out under what conditions idealization and successive concretization is a form of truth approximation. It was illustrated by the paradigm of concretization: the transition of the ideal gas law to the law of Van der Waals. I like to mention here that I demonstrated in Kuipers [1985] that the kinetic explanation of the law of Van der Waals can be reconstrued as a concretization of the kinetic explanation of the ideal gas law. In the present article, it was finally shown that concretization is also an important strategy in validity research for interesting theorems.

This article clearly illustrates that idealization and concretization is also a useful strategy in philosophy. At least for myself, I am sure that I would never have found the refined explication, without first developing the naive one.

It will be no surprise that the conceptual and methodological features of the refined explication are concretizations of features of the naive one. A number of such feature-concretizations are presented in Kuipers [1990]. Among others it is shown there that my argument in Kuipers [1989] for the explanation of the global success of (natural) science on the basis of naive truthlikeness can be concretized to an argument based on refined truthlikeness.

Moreover, from that article it becomes clear that the notion of refined truthlikeness can relativize all kinds of incommensurability claims between theories formulated within different conceptual frames. In particular it is demonstrated that refined truthlikeness of stratified theories, containing theoretical terms, is under general conditions projectable on observational theories. By consequence, refined truthlikeness paves the way for fundamental, or at least pragmatic, commensurability of related theories.

Department of Philosophy
University of Groningen
Aweg 30
9718 CW Groningen
The Netherlands

* I like to acknowledge that van Benthem [1987] played in several respects a crucial role in the research for a new refined definition. Moreover, I like to thank David Miller and Ilkka Niiniluoto for their comments on an earlier version.

REFERENCES

Benthem, J. van [1987]. Verisimilitude and conditionals. In: Kuipers [1987], pp. 103-128.

Hamminga, B. [1983]. *Neoclassical Theory Structure and Theory Development*. Berlin: Springer.

Krajewski, W. [1977]. *Correspondence Principle and Growth of Science*. Dordrecht: Reidel.

Kuipers, T. [1982]. Approaching descriptive and theoretical truth. *Erkenntnis*, **18**, 343-387.

Kuipers, T. [1984]. Approaching the truth with the rule of success. *Philosophia Naturalis*, **21**, 244-253.

Kuipers, T. [1985]. The paradigm of concretization: the Law of Van der Waals. In: J. Brzeziński (Ed.), *Consciousness: Methodological and Psychological Approaches.* (*Poznań Studies in the Philosophy of the Sciences and the Humanities*, **8**). Amsterdam: Rodopi, pp. 185-199.

Kuipers, T. [1987]. A structuralist approach to truthlikeness. In: Kuipers, T.A.F. [1987] (Ed.). *What is closer-to-the-truth?* (*Poznań Studies in the Philosophy of the Sciences and the Humanities*, **10**). Amsterdam: Rodopi, pp. 79-99.

Kuipers, T. [1989]. How to explain the success of the natural sciences. *Proceeding of the 13th Wittgenstein Symposium*. Vienna, pp. 318-322.

Kuipers, T. [1990]. Naive and refined truth approximation (to appear in *Synthese*).

Niiniluoto, I. [1986]. Theories, approximations, and idealizations. In: R. Barcan Marcus *et al.* (Eds.) *Logic, Methodology and Philosophy of Science*, **7**. Amsterdam: North-Holland, pp. 255-289 (revised and extended version in: J. Brzeziński, F. Coniglione, T.A.F. Kuipers and L. Nowak [1990] (Eds.). *Consciousness: Idealization I: General Problems*. (*Poznań Studies in the Philosophy of the Sciences and the Humanities*, **16**). Amsterdam-Atlanta, GA: Rodopi, pp. 9-57.

Nowak, L. [1980]. *The Structure of Idealization*. Dordrecht: Reidel.